

Copyright (c) 1993 - 2000 Compugen Ltd.	GenCore - version 4.5	
Run on: March 1, 2001, 16:20:32 ; Search time 17.21 seconds (without alignments)	Search time 17.21 seconds (without alignments)	
Sequence: US-09-331-631a-24_COPY_29_94	123.847 Million cell updates/sec	
Scoring table: BLOSUM62		
Perfect score: 382		
Sequence: 1 HDDEDRRGHSLQCVQRC.....EQEEQQGRGRGWHGEGERE 66		
Scoring table: Gpop 10.0 , Gapext 0.5		
Searched: 88757 seqs, 32294092 residues		
Total number of hits satisfying chosen parameters: 88757		
Minimum DB seq length: 0		
Maximum DB seq length: 200000000		
Post-processing: Minimum Match 0%		
Maximum Match 100%		
Database : SwissProt_39.*		
Pred. No is the number of result's predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		
SUMMARIES		
Result No.	Query Match Length DB ID	Description
1	133 34.8 573 1 GLBL_MAIZE	P15590 zea mays (m
2	90 23.6 588 1 VCLB_GOSHI	P09801 gossypium h
3	85 22.3 639 1 GLCX_SOYBN	P18227 glycine max
4	83.5 21.9 605 1 VCLX_GOSHI	P09799 gossypium h
5	76.5 20.0 338 1 LEGB_PEA	P14594 pisum sativ
6	73 19.1 605 1 GLCA_SOYBN	P13116 glycine max
7	72 18.8 852 1 SRCH_RABIT	P16233 oryctolagus
8	70.5 18.5 185 1 T2_MOUSE	Q06666 mus musculus
9	70 18.3 1130 1 REPT_MOUSE	P97747 mus musculus
10	70 18.3 1124 1 PER_DROME	P016624 drosophila
11	69.5 18.2 1549 1 TMRY_SHEEP	P22793 ovis aries
12	69.5 18.2 336 1 FILA_MOUSE	P11088 mus musculus
13	69.5 18.2 1208 1 PER_DRYO	P24767 drosophila
14	69 18.1 661 1 PER_DROSE	Q03354 drosophila
15	68.5 17.9 654 1 CTK4_MOUSE	Q61423 mus musculus
16	68.5 17.9 655 1 CTK4_RAT	P153883 rattus norvegicus
17	68 17.8 654 1 CTK4_MUSPP	Q28527 mustela put
18	67.5 17.7 1394 1 CNG4_BOVIN	P28187 bos taurus
19	67 17.5 228 1 DH25_ORYSA	P30287 oryza sativa
20	67 17.5 516 1 LEGB_GOSHI	P09800 gossypium h
21	67 17.5 653 1 CTK4_HUMAN	P22459 homo sapiens
22	67 17.5 1121 1 DDX8_ARATH	Q31953 arabidopsis
23	66.5 17.4 656 1 PER_DROS1	Q03355 drosophila
24	66.5 17.4 676 1 PER_DROMA	P03353 drosophila
25	66.5 17.3 407 1 IE68_HSVA	Q01042 herpesvirus
26	66 17.3 1407 1 TMRY_RABIT	P37709 oryctolagus
27	65.5 17.1 412 1 YNPL_CAEEL	P31554 caenorhabditis
28	65.5 17.1 1898 1 TMRY_HUMAN	Q07283 homo sapiens
29	65 17.0 355 1 SRI_MOSSP	Q62563 mus spretus
30	65 17.0 524 1 SBR_SOYBN	Q006762 glycine max
31	65 17.0 888 1 YGB4_YEAST	P23339 saccharomyces
32	65 17.0 1063 1 SPT5_YEAST	P27692 saccharomyces
33	17.0 2248 1 CYAL_DROME	P32870 drosophila
ALIGNMENTS		
Query Match	34.8%; Score 133; DB 1; Length 573;	

Best Local Similarity		36.2%		Pred. No. 8.5e-07;		Matches 25; Conservative 15; Mismatches 23;		Indels 6; Gaps 3;	
Qy	2	DDEDDRGGISLQOCVORCQERPRYSHARCVQECRDDQ---QDQHGRHSEEEGGRGGS 57				01-OCT-1989 (Rel. 12, Created)	DT	01-OCT-1989 (Rel. 12, Last sequence update)	DT
Db	26	DDNHMHGKGSKSGRC-EDRPHQPRCQCREERERKQRSRHEADDRSGEGS 84				01-MAR-1989 (Rel. 10, Last sequence update)	DT	01-AUG-1992 (Rel. 23, Last annotation update)	DE
Qy	58	WHGEREREE 66				15-JUL-1999 (Rel. 38, Last annotation update)	DT	BEPA-CONGLYCININ, ALPHA' CHAIN PRECURSOR.	GN
Db	85	-EDEREREQ 92						CG-1.	
RESULT 2		VCLB_GOSH1		STANDARD;		PRT; 588 AA.		Glycine max (Soybean).	
VCLB_GOSH1		PRT;		588 AA.		OC		Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;	
ID	PO9801;	PRT;		588 AA.		OC		Fabales; Fabaceae; Papilionoideae; Glycine.	
DT	01-MAR-1989 (Rel. 10, Last sequence update)	PRT;		588 AA.		RN		[1]	
DT	01-MAR-1989 (Rel. 10, Last sequence update)	PRT;		588 AA.		RN		SEQUENCE FROM N.A.	
DT	15-JUL-1999 (Rel. 38, Last annotation update)	PRT;		588 AA.		RN		MEDLINE=86210867; PubMed=3013879;	
DE	VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B).	PRT;		588 AA.		RN		Doyle J. B., Schuler M.A., Godette W.D., Zenger V., Beachy R.N.,	
OS	Gossypium hirsutum (Upland cotton)	PRT;		588 AA.		RN		Slightom J.L.;	
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	PRT;		588 AA.		RA		"The glycosylated seed storage proteins of Glycine max and Phaseolus	
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;	PRT;		588 AA.		RA		vulgaris. Structural homologies of genes and proteins."	
RL	J. Biol. Chem. 261:9228-9238(1986).	PRT;		588 AA.		RN		[2]	
RN	SEQUENCE OF 340-639 FROM N.A.	PRT;		588 AA.		RN		SEQUENCE OF 340-639 FROM N.A.	
RX	MEDLINE=83143289; PubMed=6897678;	PRT;		588 AA.		RN		Schuler M.A., Ladin B.P., Polacco J.C., Freyer G., Beachy R.N.;	
RA	STRUCTURAL SEQUENCES ARE CONSERVED IN THE GENES CODING FOR THE	PRT;		588 AA.		RA		alpha-, alpha' and beta-subunits of the soybean 7S seed storage	
RT	PROTEIN."	PRT;		588 AA.		RN		protein."	
RL	Nucleic Acids Res. 10:8245-8261(1982).	PRT;		588 AA.		RN		NUCLEIC ACIDS RES. 10:8245-8261(1982).	
CC	-I- FUNCTION: SEED STORAGE PROTEIN. ACCUMULATES DURING SEED	PRT;		588 AA.		RN		-I- FUNCTION: SEED STORAGE PROTEIN. ACCUMULATES DURING SEED	
CC	DEVELOPMENT AND IS HYDROLYZED AFTER GERMINATION TO PROVIDE A	PRT;		588 AA.		RN		DEVELOPMENT AND ITS HYDROLYZED AFTER GERMINATION TO PROVIDE A	
CC	CARBON AND NITROGEN SOURCE FOR THE DEVELOPING SEEDLING.	PRT;		588 AA.		RN		CARBON AND NITROGEN SOURCE FOR THE DEVELOPING SEEDLING.	
CC	-I- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,	PRT;		588 AA.		RN		-I- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,	
CC	CONVICILIN, CONGLYCININ, ETC.).	PRT;		588 AA.		RN		CONVICILIN, CONGLYCININ, ETC.).	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	PRT;		588 AA.		RN		This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	PRT;		588 AA.		RN		between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	PRT;		588 AA.		RN		the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	PRT;		588 AA.		RN		use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	PRT;		588 AA.		RN		modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	PRT;		588 AA.		RN		entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	PRT;		588 AA.		RN		or send an email to license@isb-sib.ch).	
DR	EMBL: M16891; AAA33071.1; -.	PRT;		588 AA.		RN		DR EMBL: M13739; AAB01374.1; -.	
DR	EMBL: J01300; ; NOT_ANNOTATED_CDS.	PRT;		588 AA.		RN		DR EMBL: J01300; ; NOT_ANNOTATED_CDS.	
DR	PIR: A30038; FWCNAB.	PRT;		588 AA.		RN		PIR: B24810; B24810.	
DR	HSSP: P50477; ICAX.	PRT;		588 AA.		RN		HSSP: P02833; 2PHL.	
DR	INTERPRO: IPR001113; -.	PRT;		588 AA.		RN		INTERPRO: IPR001113; -.	
DR	PFAM: PF00546; Seedstore_7s; 1.	PRT;		588 AA.		RN		PFAM: PF00546; Seedstore_7s; 2.	
DR	KW	Seed storage protein; Signal; Glycoprotein; Multigene family.		PRT;		588 AA.		KW	
FT	SIGNAL	1		22		PRT;		FT SIGNAL 1 22	
FT	PROPEP	23		62		PRT;		FT PROPEP 23 62	
FT	CHAIN	63		639		PRT;		FT CHAIN 63 639	
FT	CARBOHYD	277		N-LINKED (GAGNAC. . .) (POSSIBLY).		PRT;		FT CARBOHYD 277 N-LINKED (GAGNAC. . .) (POSSIBLY).	
FT	CARBOHYD	551		551		PRT;		FT CARBOHYD 551 551	
FT	CONFLICT	543		543		PRT;		FT CONFLICT 543 543 P -> L (IN REF. 2).	
FT	CONFLICT	549		549		PRT;		FT CONFLICT 549 549 M -> V (IN REF. 2).	
FT	CONFLICT	608		608		PRT;		FT CONFLICT 608 608 M -> T (IN REF. 2).	
SQ	SEQUENCE	639		639		PRT;		SQ SEQUENCE 639 639 AA; 74325 MW; 469BF24C79651E3F CRC64;	
Qy	Query Match	22.3%		Score 85;		DB 1;		Query Match 22.3% Score 85;	
Qy	Best Local Similarity	24.3%		Length 639;		DB 1;		Best Local Similarity 24.3% Length 639;	
Qy	Mismatches	20;		Conservative		9;		Mismatches 20; Conservative 9;	
Qy	Indels	2;		Gaps		52;		Indels 2; Gaps 52;	
Qy	DB	174		EQEQ 177		PRT;		DB 174 EQEQ 177	
RESULT	3	GLCX_SOYBN		STANDARD;		PRT;		RESULT 3 GLCX_SOYBN STANDARD;	
ID	GLCX_SOYBN	PRT;		639 AA.		PRT;		ID GLCX_SOYBN P11827;	
Qy	48	QEE---		-----BGRGRGWH-----GEGERE 66		Qy 48 QEE--------BGRGRGWH-----GEGERE 66		Qy 48 QEE--------BGRGRGWH-----GEGERE 66	
Db	94	EDEGEOPRPFPPRQPQPHOEEHEKEEHWKRKEKGKGSIEE 140		Db 94 EDEGEOPRPFPPRQPQPHOEEHEKEEHWKRKEKGKGSIEE 140		Db 94 EDEGEOPRPFPPRQPQPHOEEHEKEEHWKRKEKGKGSIEE 140		Db 94 EDEGEOPRPFPPRQPQPHOEEHEKEEHWKRKEKGKGSIEE 140	

CC VACUOLAR PROTEIN BODIES. SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONGLICININ, ETC.).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X17698; CAA35691.1; -.

CC PIR; S14681; FWSB.

CC HSSP; P50477; ICAW.

DR INTERPRO; IPR001113; -.

DR PROSITE; PS00328; HCP; 10.

DR Calcium-binding; Signal; Repeat.

FT SIGNAL 1 22

FT CHAIN 28 852

FT REPEAT 59 100

FT REPEAT 59 79

FT REPEAT 80 100

FT REPEAT 199 470

FT REPEAT 199 224

FT REPEAT 225 253

FT REPEAT 254 282

FT REPEAT 283 310

FT REPEAT 311 339

FT REPEAT 340 367

FT REPEAT 368 395

FT REPEAT 396 423

FT REPEAT 424 451

FT REPEAT 452 470

FT REPEAT 471 585

FT REPEAT 721 733

FT REPEAT 780 826

FT REPEAT 852 AA; 96117 MW; CBERB30506BBCB57 CRC64;

FT REPEAT 452 470

FT REPEAT 471 585

FT REPEAT 721 733

FT REPEAT 780 826

FT REPEAT 852 AA; 96117 MW; A43CCB3E94B930 CRC64;

Query Match 19.1%; Score 73; DB 1; Length 605; Best Local Similarity 30.3%; Pred. No. 1.3; Matches 20; Conservative 6; Mismatches 14; Indels 26; Gaps 3; QY 15 QCVORQEROERRY---SHARC-----VQGC-----RDQQQQHICRHEQ 48

Db 37 KCLQCSNQSERDYSRQNQACHARCNLKVKEKEEEBIPRPRPRPQHPERPQQPEKEED 96

QY 49 EEEQGR 54

Db 97 EDEQPR 102

RESULT 7

ID SRCH_RABIT STANDARD; PRT; 852 AA.

AC P16230; 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 15-JUL-1999 (Rel. 39, Last annotation update)

DE SARCOPLASMIC RETICULUM HISTIDINE-RICH CALCIUM-BINDING PROTEIN

DE PRECURSOR (HCP).

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

[1] RN 49

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=SKELETAL MUSCLE;

RX MEDLINE=9036884; PubMed=2808355;

RA Hofmann S.L., Goldstein J.L., Orth K., Moomaw C.R., Slaughter C.A., Brown M.S.,

RA RT "Molecular cloning of a histidine-rich Ca²⁺-binding protein of sarcoplasmic reticulum that contains highly conserved repeated elements.",

RL J. Biol. Chem. 264:18083-18090(1989).

CC -!- FUNCTION: HCP MAY PLAY A ROLE IN THE REGULATION OF CA(2+)

CC SEQUESTRATION OR RELEASE IN THE SR OF SKELETAL AND CARDIAC

CC MUSCLE.

CC -!- SUBCELLULAR LOCATION: SARCOPLASMIC RETICULUM LUMEN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

CC EMBL; J05080; AA31279.1; -.

CC PIR; A34373; A34373.

CC INTERPRO; IPR00134; -.

CC PROSITE; PS00328; HCP; 10.

CC Calcium-binding; Signal; Repeat.

FT SIGNAL 1 22

FT CHAIN 28 852

FT REPEAT 59 100

FT REPEAT 59 79

FT REPEAT 80 100

FT REPEAT 199 470

FT REPEAT 199 224

FT REPEAT 225 253

FT REPEAT 254 282

FT REPEAT 283 310

FT REPEAT 311 339

FT REPEAT 340 367

FT REPEAT 368 395

FT REPEAT 396 423

FT REPEAT 424 451

FT REPEAT 452 470

FT REPEAT 471 585

FT REPEAT 721 733

FT REPEAT 780 826

FT REPEAT 852 AA; 96117 MW; A43CCB3E94B930 CRC64;

FT REPEAT 452 470

FT REPEAT 471 585

FT REPEAT 721 733

FT REPEAT 780 826

FT REPEAT 852 AA; 96117 MW; A43CCB3E94B930 CRC64;

Query Match 18.8%; Score 72; DB 1; Length 852; Best Local Similarity 27.8%; Pred. No. 2.3; Matches 20; Conservative 11; Mismatches 21; Indels 20; Gaps 3; QY 7 RRGHSLQOCVQRCRERPRYSHARCVQECRQDQQHGR ---HEPEEGQRGRG---- 57

Db 663 KRGGE-----EDEEEQKOTHHSLEDEDDEEFGHRSLSQEDQEEEDRRGESAKVQA 714

QY 58 ---WHGEGEREE 66

Db 715 PLRHHHEEEEEE 726

RESULT 8

ID T2_MOUSE STANDARD; PRT; 185 AA.

AC Q00665; 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE OCTAPETIDE-REPEAT PROTEIN T2.

GN SRST OR T2.

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

[1] RN 1

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=MACROPHAGE;

RX MEDLINE=93092084; PubMed=1458435;

RA di Carlo M., Montana G., Romancino D.P., Monteleone D.;

RA "A mouse repeat sequence conserved in eukaryotic genomes.",

RL J. Submicrosc. Cytol. Pathol. 24:467-472(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X67863; CAA48048.1; -.

RL Proc. Natl. Acad. Sci. U.S.A. 91:2260-2264 (1994).
 CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN PER DOSE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE IN LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECOLOSTON BEHAVIOR, AND FOR THE RHYTHMIC COMPONENT OF THE MALE COURSHIP SONG THAT ORIGINATES IN THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CLOCK DEPENDS ON THE RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETEROODIMER COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT TRANSCRIPTIONAL INHIBITION.
 CC -!- SUBUNIT: FORMS A HETEROODIMER WITH TIMELESS (TIM); THE COMPLEX THEN TRANSLOCATES INTO THE NUCLEUS.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY. FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED FOR NUCLEAR LOCALIZATION.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST SIX ISOFORMS; PER-A/SHORT (SHOWN HERE), PER-B, PER-C, PER-D AND PER-E; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURAL TISSUES AND IN SEVERAL NONNEURAL TISSUES OF THE ABDOMEN. MALPIGHTIAN TUBULES CONTAIN A CIRCADIAN PACEMAKER THAT FUNCTIONS INDEPENDENTLY OF THE BRAIN. EXPRESSION OSCILLATES IN ALL TISSUES STUDIED EXCEPT FOR THE OVARY.
 CC -!- ALTERNATIVE ISOFORMS: MAINLY EXPRESSED IN ADULT'S HEAD.
 CC -!- INDUCTION: EXPRESSION IS SENSITIVE TO TEMPERATURE BUT NOT TO LIGHT.
 CC -!- DOMAIN: CONTAINS A REMARKABLE RUN OF ALTERNATING GLY-THR RESIDUES WHICH IS POLYMORPHIC IN LENGTH. AT LEAST THREE TYPES OF GLY-THR LENGTH EXIST IN THE NATURAL POPULATION, (GLY-THR)23 (SHOWN HERE), AND TWO MAJOR VARIANTS (GLY-THR)17 AND (GLY-THR)20. THIS GLY-THR STRETCH IS IMPLICATED IN THE MAINTENANCE OF CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT LEADS TO A SHORTENING OF THE COURSHIP SONG CYCLE PERIOD AND THUS COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC MATING BEHAVIOR.
 CC -!- DOMAIN: MUTATIONS IN THE PAS DOMAIN RESULT IN LONGER CIRCADIAN RHYTHMS AND COURSHIP SONG (PERL MUTATION) OR MAKES THE FLIES ARRHYTHMIC (PERL MUTATION).
 CC -!- PTM: PHOSPHORYLATION WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE DOUBLE-TIME PROTEIN (DFT). PHOSPHORYLATION COULD BE IMPLICATED IN THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETEROODIMER PER-TM.
 CC -!- SIMILARITY: TO OTHER G-T STRETCH CONTAINING PROTEINS.
 CC -!- SIMILARITY: CONTAINS A PAS (PER-ARMY-SIM) DIMERIZATION DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query Match 18.3%; score 70; DB 1; Length 1224;
 Best Local Similarity 36.24; Pred. No. 5.3; Matches 21; Conservative 9; Mismatches 14; Indels 14; Gaps 4;
 CC -!- EMBL: M30114; AAA28752.1; -.
 CC -!- EMBL: M30114; AAA28754.1; -.
 CC -!- EMBL: AF033029; AAB8476.1; -.
 CC -!- EMBL: X03636; CAA27585.1; -.
 CC -!- EMBL: M11069; AAA28751.1; -.
 CC -!- EMBL: AL024485; CAA19677.1; -.
 CC -!- EMBL: AL024485; CAA19678.1; -.
 CC -!- EMBL: AL024485; CAA19679.1; -.
 CC -!- EMBL: AL024485; CAA19680.1; -.
 CC -!- EMBL: M13355; AAA28785.1; -.
 CC -!- EMBL: D00009; BAA00007.1; -.
 CC -!- EMBL: L07817; AAA28777.1; -.
 CC -!- EMBL: L07818; AAA28776.1; -.
 CC -!- EMBL: L07819; AAA28775.1; -.
 DR EMBL: L07821; AAA28773.1; -.
 DR EMBL: L07823; AAA28771.1; -.
 DR EMBL: L07825; AAA28769.1; -.
 DR EMBL: A23932; UMPF.
 DR FLYBASE; FBgn003068; per.
 DR INTERPRO; IPR000014; -.
 DR PFAM; PF00989; PAS; 2.
 DR KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation; Polymorphism; Alternative splicing.
 DR KW NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 66 79
 FT REPEAT 227 295
 FT REPEAT 378 445
 FT DOMAIN 459 499
 FT DOMAIN 694 743
 FT DOMAIN 74 79
 FT DOMAIN 143 146
 FT DOMAIN 223 226
 FT DOMAIN 872 879
 FT DOMAIN 898 907
 FT DOMAIN 906 914
 FT DOMAIN 1006 1013
 FT DOMAIN 1035 1041
 FT DOMAIN 749 868
 FT VARSPLIC 1 62
 FT VARSPLIC 868 963
 FT VARSPLIC 863 958
 FT VARSPLIC 1076 1224
 FT VARSPLIC 1155 1224
 FT VARIANT 697 702
 FT VARIANT 697 708
 FT CONFLICT 211 211
 FT CONFLICT 498 499
 FT CONFLICT 637 637
 FT CONFLICT 762 762
 FT CONFLICT 762 764
 FT CONFLICT 1029 1029
 FT CONFLICT 1038 1038
 FT CONFLICT 1075 1075
 FT CONFLICT 1114 1114
 FT CONFLICT 1215 1215
 FT CONFLICT 1224 AA; 1278552 MN; 71FF54ECF3E50F4A CRC64;
 SQ SEQUENCE

RESULT 11
 TRHY_SHEEP TRHY_SHEEP STANDARD; PRT; 1549 AA.
 AC P22793;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TRICHOHYALIN.
 GN THH.
 OS Ovis aries (Sheep).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

RP SEQUENCE FROM N.A.
 RX MEDLINE:9326018; PubMed:7684041;
 RA Flett M.J., McLaughlin C.J., Campbell M.T., Rogers G.E.;
 RT "Analysis of the sheep trichohyalin gene: potential structural and
 RT calcium-binding roles of trichohyalin in the hair follicle";
 RL J. Cell Biol. 121: 855-865 (1993).
 RN [2]
 RP SEQUENCE OF 1016-1549 FROM N.A.
 RC STRAIN=MERINO-DORSET HORN X BORDER LEICESTER; TISSUE=WOOL FOLLICLES;
 RX MEDLINE=90103632; PubMed=2298812;
 RA Flett M.J., Prestland R.B., Rogers G.E.;
 RT "The cDNA-derived amino acid sequence for trichohyalin, a
 RT differentiation marker in the hair follicle, contains a 23 amino acid
 RT repeat." J. Cell Biol. 110: 427-436 (1990).
 CC -1- FUNCTION: INTERMEDIATE FILAMENT-ASSOCIATED PROTEIN THAT ASSOCIATES
 IN REGULAR ARRAYS WITH KERATIN INTERMEDIATE FILAMENTS (KIF) OF THE
 INNER ROOT SHEATH CELLS OF THE HAIR FOLLICLE AND THE GRANULAR
 LAYER OF THE EPIDERMIS. IT LATER BECOMES CROSS-LINKED TO KIF BY
 ISODIPEPTIDE BONDS. IT MAY SERVE AS SCAFFOLD PROTEIN, TOGETHER
 WITH INVOLUCRIN, IN THE ORGANIZATION OF THE CELL ENVELOPE OR EVEN
 ANCHOR THE CELL ENVELOPE TO THE KIF NETWORK. IT MAY BE INVOLVED IN
 ITS OWN CALCIUM-DEPENDENT POSTSYNTHETIC PROCESSING DURING TERMINAL
 DIFFERENTIATION.
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST TWO FORMS MAY BE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE HARD KERATINIZING TISSUES SUCH AS
 THE INNER ROOT SHEATH (IRS) OF HAIR FOLLICLES AND MEDULLA, AND IN
 THE EPITHELIA OF THE TONGUE, HOOF AND RUMEN.
 CC -1- DOMAIN: CONSISTS OF NINE DOMAINS. DOMAIN 1 CONTAINS TWO EF-HAND
 CALCIUM BINDING DOMAINS. DOMAINS 2, 4, 6, AND 8 ARE ALMOST
 ENTIRELY ALPHA-HELICAL, CONFIGURED AS A SERIES OF PEPTIDE REPEATS
 OF VARYING REGULARITY, AND ARE THOUGHT TO FORM A SINGLE-STRANDED
 ALPHA HELICAL ROD STABILIZED BY IONIC INTERACTIONS. DOMAIN 6 IS
 THE MOST REGULAR AND MAY BIND KIF DIRECTLY BY IONIC INTERACTIONS.
 CC DOMAINS 5 AND 7 ARE LESS WELL ORGANIZED AND MAY INDUCE FOLDS IN
 THE MOLECULE. DOMAIN 9 CONTAINS THE C-TERMINUS, CONSERVED AMONG
 CC DIFFERENT SPECIES.
 CC -1- PTM: KNOWN SUBSTRATE OF TRANSGLUTAMINASE. SOME 200 ARGININES ARE
 CC PROBABLY CONVERTED TO CITRULLINES BY PEPTIDYLARGININE DEIMIDASE.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE S-100
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 CC purposes requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z18361; CAA79165.1; -.
 DR EMBL; X51695; CAA35992.1; -.
 DR PIR; A34209; A34209.
 DR PIR; S32633; S32633.
 DR PIR; A40691; A40691.
 DR HSSP; P02633; 31CB.
 DR INTERPRO; IPR010751; -.
 DR INTERPRO; IPR020208; -.
 DR PFAM; PF01023; S100; 1.
 DR PFAM; PF00036; effhand; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; FALSE_NEG.
 KW Repeat; Calcium-binding.
 FT DOMAIN 1 91 S 100 LIKE.
 FT CA_BIND 22 33 SITE I (LOW AFFINITY) (POTENTIAL).
 FT CA_BIND 62 73 SITE II (HIGH AFFINITY) (POTENTIAL).
 FT DOMAIN 413 832 14 X 28 AA APPROXIMATE TANDEM REPEATS.

Query	Match	Score	DB	Length
QY	18: 3%; Best Local Similarity Matches 23; Conservative	70;	DB 1;	Length 1549;
Db	3 DEDDRGGHSLQQVCQRCRQERPRYSHARQVQECCRDDQ--QHGRHEQEQQGRGRGWIQ 1454 EKKRRQEQLRQEQRROERER-KFREEEQLRQQQEKKRQERQWVQOSRRQWIE 1511	4;	Gaps 2;	
		10;	Mismatches 1;	
		29;	Indels 4;	
		1-1.		
FT	REPEAT 413	448		
FT	REPEAT 449	476		
FT	REPEAT 477	504	1-3.	
FT	REPEAT 505	532	1-4.	
FT	REPEAT 533	560	1-5.	
FT	REPEAT 561	588	1-6.	
FT	REPEAT 589	616	1-7.	
FT	REPEAT 617	644	1-8.	
FT	REPEAT 645	678	1-9.	
FT	REPEAT 679	706	1-10.	
FT	REPEAT 707	742	1-11.	
FT	REPEAT 743	771	1-12.	
FT	REPEAT 772	796	1-13.	
FT	REPEAT 797	832	1-14.	
FT	REPEAT 938	1507	23 x 23 AA APPROXIMATE TANDEM REPEATS.	
FT	REPEAT 938	961	2-1.	
FT	REPEAT 962	985	2-2.	
FT	REPEAT 986	1021	2-3.	
FT	REPEAT 1022	1044	2-4.	
FT	REPEAT 1045	1067	2-5.	
FT	REPEAT 1068	1090	2-6.	
FT	REPEAT 1091	1121	2-7.	
FT	REPEAT 1122	1144	2-8.	
FT	REPEAT 1145	1167	2-9.	
FT	REPEAT 1168	1197	2-10.	
FT	REPEAT 1198	1227	2-11.	
FT	REPEAT 1228	1250	2-12.	
FT	REPEAT 1251	1273	2-13.	
FT	REPEAT 1274	1296	2-14.	
FT	REPEAT 1297	1319	2-15.	
FT	REPEAT 1320	1342	2-16.	
FT	REPEAT 1343	1368	2-17.	
FT	REPEAT 1369	1391	2-18.	
FT	REPEAT 1392	1416	2-19.	
FT	REPEAT 1417	1439	2-20.	
FT	REPEAT 1440	1461	2-21.	
FT	REPEAT 1462	1484	2-22.	
FT	REPEAT 1485	1507	2-23.	
FT	VARIANT 1145	1197	MISSING (IN SHORT FORM).	
FT	VARIANT 1251	1273	MISSING (IN SHORT FORM).	
FT	CONFECT 1399	1399	E -> G (IN REF. 2);	
SQ	SEQUENCE	1549 AA;	201173 MW;	E72FB9FF1326E54E CRC64;

PER_DROSE
ID PER_DROSE STANDARD; PRT; 661 AA.
AC 00354; 026285;
DT 01-OCT-1993 (Rel. 27, created)
15-JUL-1998 (Rel. 38, last sequence update)
30-MAY-2000 (Rel. 39, last annotation update)
DE PERIOD CIRCADIAN PROTEIN (FRAGMENTS).
GN PER.
OS Drosophilidae; Drosophila.
OC Drosophilidae; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydriidae; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE OF 1-558 FROM N.A.;
RX MEDLINE:93110641; PubMed=84366278;
RA Klinman R.M., Hey J.;
RT "DNA sequence variation at the period locus within and among species of the Drosophila melanogaster complex.",
Genetics 133:375-387 (1993).
RN [2]
RP SEQUENCE OF 559-661 FROM N.A.;
MEDLINE:93140158; PubMed=1407825;
RA Peixoto A.A., Costa R., Wheeler D.A., Hall J.C., Kyriacou C.P.;
RT "Evolution of the threonine/glycine repeat region of the period gene in the melanogaster species subgroup of Drosophila.";
J. Mol. Evol. 35:411-419 (1992).
RL J. Mol. Evol. 35:411-419 (1992).
CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE PERIOD LENGTH OF CIRCADIAN AND ULTRADAY RHYTHMS; AN INCREASE IN PER DOSE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERO-DIMER COORDINATELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
CC -!- SUBUNIT: FORMS HETERO-DIMER WITH TIMELESS (TIM); THE COMPLEX THEN TRANLOCATES INTO THE NUCLEUS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY. FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
CC -!- DOMAIN: THE RUN OF GLY-THR IS IMPLICATED IN THE MAINTENANCE OF CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT LEADS TO A SHORTENING OF THE COURSHIP SONG CYCLE PERIOD, AND THUS COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC MATING BEHAVIOR (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATION (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERO-DIMER PER-TIM (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS A PASS (PER-ARNT-SIM) DIMERIZATION DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; L07820; AAB2874_1; -.
DR EMBL; L07822; AAB2872_1; -.
DR EMBL; L07824; AAB28770_1; -.
DR EMBL; L07827; AAB28767_1; -.
DR EMBL; S23297; AAB25028_1; -.
DR FLYBASE; FBgn0012799; DsecPer.
DR PFAM; PF00989; PAS; 2.
KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.

RESULT 15
CIRK4_MOUSE
ID CIRK4_MOUSE STANDARD; PRT; 654 AA.
AC 061423;
DT 15-JUL-1998 (Rel. 36, created)
15-JUL-1998 (Rel. 36, last sequence update)
15-JUL-1998 (Rel. 36, last annotation update)
DE VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.4.
GN KCNA4.
OS Mus musculus (Mouse).
OC Eutheria; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NKR;
RX MEDLINE:9429198; PubMed=8020965;
RA Wyman R.S., Korenberg J.R., Kinosita K.D., Aiyar J., Coyne C.,
Chen X.N., Hustad C.M., Copeland N.G., Gutman G.A., Jenkins N.A.,
Chandy K.G.;
RT "Genomic organization, nucleotide sequence, biophysical properties, and localization of the voltage-gated K⁺ channel gene KCNA4/Kv1.4 to mouse chromosome 2/human 11p14.3-p15.2 and mapping of KCNC1/Kv3.1 to mouse 7/human 11p14.3-p15.2 and KCNA1/Kv1.1 to human 12p13.1.";
RT Genomics 20:191-202 (1994).
RL -
CC -!- FUNCTION: THIS PROTEIN MEDIATES THE VOLTAGE-DEPENDENT POTASSIUM ION PERMEABILITY OF EXCITABLE MEMBRANES. ASSUMING OPENED OR CLOSED CONFORMATION IN RESPONSE TO THE VOLTAGE DIFFERENCE ACROSS THE MEMBRANE, THE PROTEIN FORMS A POTASSIUM-SELECTIVE CHANNEL. THROUGH WHICH K⁺ IONS MAY PASS IN ACCORDANCE WITH THEIR ELECTROCHEMICAL GRADIENT.
CC -!- SUBUNIT: HETERO-TETRAMER OF POTASSIUM CHANNEL PROTEINS (PROBABLE).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- DOMAIN: THE AMINO TERMINUS MAY BE IMPORTANT IN DETERMINING THE RATE OF INACTIVATION OF THE CHANNEL WHILE THE TAIL MAY PLAY A ROLE IN MODULATION OF CHANNEL ACTIVITY AND/OR TARGETING OF THE CHANNEL TO SPECIFIC SUBCELLULAR COMPARTMENTS.
CC -!- DOMAIN: THE SEGMENT S4 IS PROBABLY THE VOLTAGE-SENSOR AND IS CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- SIMILARITY: THIS CHANNEL PROTEIN BELONGS TO THE A-TYPE POTASSIUM CURRENT CLASS. BELONGS TO SHAKER POTASSIUM CHANNEL SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC

```

Query Match          17.9%; Score 68.5; DB 1; Length 654;
Best Local Similarity 27.3%; Pred. No. 4.1;
Matches 15; Conservative 10; Mismatches 27; Indels 3; Gaps 1
Qy   11 HSIQOCVORCQERPRYSHARCQVOCRDDQQHGRHEQPEOGR--GROWHGEG 62
      |   |   |   |   |   :   :   :   :   :   :   :   :   :   :   :   :   :
      96 HRSQSFPCSDLMPSGSEKILRELSEFEEDEEEEEEEEGRPFYVSBDHGDG 150
Db

```

Search completed: March 1, 2001, 16:20:34
Job time: 209 sec

Search completed: March 1, 2001, 16:20:34
Job time: 209 sec

Search completed: March 1, 2001, 16:20:34
Job time: 209 sec